Integral Transform Parameter Estimation

Martha Contreras* George Casella†

Cornell University

September 18, 1996

Abstract

There are many reasons for considering estimation in a transformed version of a problem. In this paper we look at a class of compartment models, and see that it is possible to estimate the underlying parameters more easily in a transformed problem. In particular, it is not necessary to know the form of the regression function in order to perform the estimation in the Laplace space. We show how to construct estimators of the underlying model coefficients that are consistent and asymptotically normal. The effectiveness of the estimation method is also illustrated with a simulation and an analysis of data.

Key words and phrases : Least squares, Integral transforms, Compartment models, Pharmacokinetics, Integral Laplace approximations, saddlepoint approximations.

*The author thanks Louise M. Ryan for her introduction to statistical transform estimation and Carlos Castillo-Chavez for valuable discussions on compartment models.

†This research was supported by NSF Grant No. DMS-9625440. This is paper BU-1363-M in the Biometrics Unit, Cornell University, Ithaca, NY 14853. File trans3.tex

1 Introduction

Estimation following transformation is a common practice in statistics, particularly in regression problems. In this paper we focus on a particular transformation–the Laplace transform–with a particular purpose. We study a class of models in which estimation of the underlying parameters is more straightforward in the Laplace space than in the original space. Although we focus on a somewhat specialized problem, our techniques can be applied to more general setups.

In the study of the dynamics of biological systems, it is common to subdivide the system under consideration into subsystems or compartments where it is possible to measure the flow of organic or inorganic materials. A central statistical problem is that of estimating from experimental data the transition rates that model the flow of material from compartment to compartment. These rates are embedded in a system of differential equations that in most cases must first be solved either analytically or numerically before the estimation can be performed. This, compounded by the fact that it might only be practical to gather experimental data for some compartments, makes accurate estimation of the rates more sensitive and costly. For example, the spectral decomposition of the compartment matrix need not exist [Moler and Van Loan (1978)].

In this paper, we propose a computationally feasible method of estimation based on integral transform methods and integral approximations applied at the differential equation level. In particular, we will see that this readily yields a closed form expression of the concentration curve at the measuring site as a function of the variable of the transformation. Through simulations and the analysis of a data set, we will see that optimizing over the integral transform space yields estimates comparable to the standard least squares estimates, and that are consistent and asymptotically normal. However, we avoid having to solve the system of differential equations, which requires knowledge of the exact form of the function, and must be done to compute the least squares solution.

Parameter estimation using integral transform methods of estimation has an extensive history dating back at least twenty years. Some of the people who have done work in this area are Schuh and Tweedie (1979), Feuerverger and McDunnough ((1981ab), (1984)), Feigin, Tweedie and Belyea (1983), Wells and Bryant (1984), Bagchi, Ord, and Sullivan (1986), and Laurence

and Morgan (1987). Using illustrations from reliability theory [Wells and Bryant (1984)], examples where the distribution of interest arises from a convolution [Bagchi, Ord, and Sullivan (1986)], and other examples where the density cannot be readily obtained (but some integral transform, like the Laplace or Fourier transform can), these authors demonstrated the practicality and advantages of transform estimation over other existing methods of estimation.

However, the advantages of transform estimation, as introduced by the above workers, were offset by necessitating that the user supply the "optimal" variable of the transformation. To address optimality in the choice of the variable of the transformation, say s, Feigin, Tweedie and Belyea (1983)suggested picking the s which yielded minimal variance over a simulated data set. Laurence and Morgan (1987)then concluded that bias was also a concern and suggested picking the variable of the transformation that yielded the minimal mean-squared error. However, any one of their estimates always depended on the choice of the variable of the transformation (and to a large degree, except for our recommendation of the s space, Ω_2, we have not solved this, either).

Like the other authors, Feuerverger and McDunnough [(1981ab), (1984)] also proposed a series of transform estimation methods which also necessitated that the user supply the variable(s) of the transformation; but unlike the others, they showed that for a finely spaced and extended sequence of variables of the transformation, their estimates were not only asymptotically normal but, that with the proper choice of the weight vector, could be made arbitrarily efficient. The present paper could be considered an extension of the work of Feuerverger and McDunnough, where we consider a more complex model, and alternate methods of dealing with the objective function (such as saddlepoint approximations). In particular, aside from the recommendation that we make on the s_j's, on the "weights" given by Δs_j, and on our choice of the approximation to $L[m](s_j)$, the objective function in (13) is one of their recommended criteria even though they did not deal with a regression model such as (5) .

In Section 2, we give a brief introduction to the theory of linear compartment models and provide an illustration of a compartment model commonly used in Pharmacokinetics, that is, the study of the movement of drugs through the body. In Section 3, we derive the transformed model by integral transforming each side of the standard least squares model. In addition, we

investigate the analytical properties of our integral transform functional, and we show that for our particular class of functions, this is always defined. We will see that this approach preserves the error structure of the original model while yielding a closed form expression for the concentration curve at the measuring site, even though it will be a function of the variable of the transformation. In Section 4, we introduce our objective function for derivation of the least squares estimate in the transformed problem, and two integral approximation schemes–the standard Riemann approximation to an integral and the Laplace integral approximation. Moreover, we establish that the resulting estimator is consistent and asymptotically normal. In Section 5, using an intravenous dose single compartment model, the numerical performance of the above two approximations, the actual objective function, and the standard least squares estimator are analyzed. This we do using both simulated data and a real data set from [Kwan, et al (1975)] in which 6 healthy male volunteers each received an initial dose of indomethacin-2^{14}-C by intravenous route. In Section 6, we give a brief summary and discussion and lastly, in the appendix, Section 7, we give the technical details of our estimation methodology.

2 Compartment Models

In this section, we provide a brief review of compartment models. The reader desiring more detail on this vast subject is referred to Jacquez (1985), Anderson (1983), or Seber and Wild (1989). Prior to introducing compartment models in generality, because it will make our presentation clearer to the reader less familiar with the topic, we consider the following simple compartment model, known in the literature as a one compartment open model with first-order absorption.

This model may be thought of as representing the single dose oral administration of a drug. Measurements of the drug's concentration are to be taken in the plasma, which is represented by the second box. The first box in this case could be representing the stomach. The flows between the compartments are denoted by the directed arrows.

$$\square \xrightarrow{\theta_1} \square \xrightarrow{\theta_2} . \tag{1}$$

A system of differential equations describing how the concentration changes

through time in (1) is given by

$$\begin{aligned} \frac{dm_1(t)}{dt} &= -\theta_1 m_1(t) \\ \frac{dm_2(t)}{dt} &= \theta_1 m_1(t) - \theta_2 m_2(t), \end{aligned} \tag{2}$$

where $m(t)$ is a vector with components $m_j(t)$, which denote the state of the system in compartment j at time t with initial input vector (transposed) $m(0) = (\xi_1, 0)^{\mathrm{T}}$. Using either linear algebraic techniques or elementary theory of linear ordinary differential equations, the solution to (2) can be seen to be

$$\begin{aligned} m_1(t) &= \xi_1 e^{-\theta_1 t} \\ m_2(t) &= \frac{\xi_1 \theta_1}{\theta_2 - \theta_1}(e^{-\theta_1 t} - e^{-\theta_2 t}). \end{aligned}$$

The objective is to estimate the constant flow rates, θ_1 and θ_2, at the measurement site, $m_2(t)$. Note once again that the targets of inference are the parameter values θ_1 and θ_2. There is less interest in estimating $m(t)$. However, to employ a standard method of estimation such as least squares, which is often used in fitting compartment models, one would necessarily need a closed form expression of the concentration function at the measurement site. This means, as the above example illustrates, that we need to solve (2).

For a general compartment model with constant flow rates, the concentration curve at any compartment is implicitly given by the system of differential equations

$$\frac{d}{dt} m(t, \theta) = A_\theta m(t, \theta) + b(t), \quad m(0, \theta) = \xi, \tag{3}$$

where θ represents the vector of nonzero flow rates from one compartment to another, $A_\theta = [\theta_{jk}]$, where j and k range between 1 and p, with p being the number of compartments, $\frac{d}{dt} m(t, \theta)$ denotes the rate of change of the material through the system for all time, and $b(t)$ is the vector whose components denote the input rate from the environment into the system. Lastly, $m(0, \theta)$ denotes the initial state of the system. Note that we now write $m(t, \theta)$ to highlight the dependence of the function m on the vector θ.

The matrix A_θ is the square compartment matrix. Such a matrix has non-negative off diagonal elements, non-positive diagonal elements, and its

column sums are non-positive. This implies that the eigenvalues of A_θ have a non-positive real part and that none are purely imaginary. [This is Gerschgorin's Theorem which can be found in Anderson (1983).] We will refer to these properties of A_θ in sections to come (It is relatively straightforward to verify that A_θ as given by (2) meets these conditions.). It is a standard result, that when A_θ is square , the solution to (3) is

$$m(t,\theta) = e^{A_\theta t}\xi + \int_o^t e^{A_\theta(t-u)}b(u)du. \tag{4}$$

Moreover, in the absence of inputs into the system, that is $b = 0$, it is known that any solution of (3) converges to an equilibrium/steady state solution as $t \to \infty$ [Hearon (1963)]. This is obvious in our illustration, (2), since if $t \to \infty$, the steady state solution gives $A_\theta m = 0$ where A_θ is nonsingular since neither θ_1 nor θ_2 are zero, thus the only solution of $A_\theta m = 0$ is $m(t) = 0$. Thus, we expect to have a zero concentration of the drug in the system for sufficiently large time.

3 The Transformed Model

Upon obtaining a closed form of the concentration curve, a standard method of estimation is least squares in the model

$$y_{ij} = m_j(t_{ij},\theta) + \epsilon_{ij}, i = 1,2,...,n_j, \quad j = 1,2,...,p, \tag{5}$$

where j ranges over the measured compartments and the ϵ_{ij} are typically *i.i.d.* normally distributed, mean zero, and have constant variance. For the remainder of this paper and without lost of generality, we take $j = 1$, that is, we only have measurements on one compartment, $m(t,\theta)$.

Examination of the model (5), especially if $m(t,\theta)$ is unknown, might suggest a nonparametric regression methodology. This route would be reasonable if there were interest in the response function, but recall that our chief interest is in estimation of the parameter vector θ, which represents the underlying flow rates. As such, nonparametric regression is not appropriate here.

To present our methodology, we first introduce a continuous analogue of (5); that is, we suppose there exist functions $y(t)$, $\epsilon(t) \in C^\infty(\Omega_1)$ where

$\Omega_1 = [0, \infty)$, $y(t_i) = y_i$ and $\epsilon(t_i) = \epsilon_i$ for all i and such that

$$y(t) = m(t, \theta) + \epsilon(t), \quad t \in \Omega_1, \quad \theta \text{ fixed}, \tag{6}$$

such that $E(\epsilon(t)) = 0$ and $Var(\epsilon(t_i)) = Var(\epsilon_i) = \sigma^2$, a positive constant. Then, upon multiplying both sides of (6) by e^{-st}, for $s > 0$, and integrating over $t \in \Omega_1$, we obtain

$$\int_{\Omega_1} e^{-st} y(t) dt = \int_{\Omega_1} e^{-st} m(t, \theta) dt + \int_{\Omega_1} e^{-st} \epsilon(t) dt, \quad s > 0. \tag{7}$$

However, since in practice we do not have either $y(t)$ or $\epsilon(t)$, but rather only observe $y(t_i) = y_i$ and $\epsilon(t_i) = \epsilon_i$, we work with the following discrete approximation to (7),

$$\sum_{i=1}^{n-1} e^{-st_i} y_i \Delta t_i \approx \sum_{i=1}^{n-1} e^{-st_i} m(t_i, \theta) \Delta t_i + \sum_{i=1}^{n-1} e^{-st_i} \epsilon_i \Delta t_i, \quad s > 0, \tag{8}$$

where $\Delta t_i = t_{i+1} - t_i$ and t_1 and t_n correspond to the initial time and the last time a measurement was made on the system, respectively.

Note that (7) and (8) are simply integral transform versions of the model (6), or integral transform approximation versions of the model (5). Furthermore, since the ϵ_i are assumed to be *i.i.d.* normal, mean zero, and to have constant variance, it is known that the errors in (8) are also *i.i.d.* normal, mean zero, and have variance which depends on s. We will use this model to estimate the parameter vector θ, for we can do so without explicit knowledge of the form of $m(t, \theta)$. To see why this is so, let us analyze the functional form

$$L[m](s, \theta) := \int_{\Omega} e^{-st} m(t, \theta) dt, \quad s > 0. \tag{9}$$

over some interval $\Omega \subseteq \Omega_1 = [0, \infty)$. If we assume that s can be complex-valued, then L in (9) is known as the Laplace transform; however, since s in this paper will be strictly real and positive, we will simply refer to L as an integral transform.

It is a fact that a closed-form solution of (9) can be obtained, without the knowledge of a closed form of $m(t, \theta)$, directly from the differential equation given by (3). One simply multiplies both sides of (3) by e^{-st} and then integrates over the time interval, Ω. Upon performing integration by parts, we obtain a closed form expression of (9) that incorporates all of the data

given by the biological process assumed to satisfy (3). That is, if $\Omega = \Omega_1$ then it follows that

$$L[m](s,\theta) = (L[b](s) + \xi)(sI - A_\theta)^{-1}. \tag{10}$$

Thus, if we could integrate over the entire positive real line, we could express the transform $L[m](s,\theta)$ in this simple form. However, in analyzing our model we would only be concerned with the range $[t_1, t_n]$, and integration by parts yields that for $\Omega = [t_1, t_n]$,

$$L[m](s,\theta) = \left\{[L[b](s) + [e^{-st_1}m(t_1,\theta) - e^{-st_n}m(t_n,\theta)]\right\}(sI - A_\theta)^{-1}.$$

This restricted range necessitates knowledge of the function $m(t,\theta)$ at the endpoints. To accommodate this we replace $m(t_1,\theta)$ and $m(t_n,\theta)$ by their empirical estimates y_1 and y_n, respectively, to obtain

$$L[m](s,\theta) = [L[b](s) + (e^{-st_1}y_1 - e^{-st_n}y_n)](sI - A_\theta)^{-1}. \tag{11}$$

Multiplying both sides of (11) by $(sI - A_\theta)$ and employing Gaussian elimination, one can see that it is possible to obtain a closed form of the concentration curve at the measurement site but in the transformed space. In particular, L applied to our illustration (2) yields

$$L[m_2](s) = \frac{\xi_1\theta_1}{(s+\theta_1)(s+\theta_2)},$$

as the transformed closed form expression of the concentration curve at the measurement site.

To address the existence of (10) or (11), for our class of functions, we establish the next claim.

Theorem 3.1 *Given that ξ is a constant vector of initial inputs and that A_θ is a compartment matrix, if $L[b](s)$ exists, it follows that (10) or similarly (11) are well-defined in Ω_1.*

Proof. Since A_θ is a compartment matrix, Gerschgorin's Theorem, Anderson (1983), establishes that its eigenvalues have non-positive real part and none are purely imaginary. This implies that $s > 0$ can never be an eigenvalue of A_θ. Therefore, $(sI - A_\theta)$ is invertible for all $s > 0$. Thus (10) and (11) exist and Theorem 3.1 is established. □

Note that there are several functions meeting the conditions of Theorem 3.1. For instance, in the absence of inputs, $b = 0$, into the system it follows that $L[b](s) = 0$, thus in this case $L[b](s)$ certainly exists. A compartment model which has $b = 0$ is said to be closed, that is, it has no inputs or outputs into the environment. Such model is referred to as a "donor controlled" model, and it is used to model, for instance, the flow of energy or nutrients through an ecosystem [Walter (1979)]. Thus our operator L is defined for this class of functions.

4 Parameter Estimation

We now turn to estimating θ in model (8) using least squares. That is, with L as given by (11), we solve the following problem

$$\min_{\theta \in \Theta} \int_{\Omega_2} [L[m(t,\theta)](s) - \sum_{i=1}^{n-1} e^{-st_i} y_i \Delta t_i]^2 ds, \tag{12}$$

where $\Omega_2 = [t_1, t_n]$ if t_1 is not identically zero. If $t_1 = 0$ then we let $\Omega_2 = [\epsilon, t_n]$ where $\epsilon > 0$ but small. This condition is needed so that the integrand in (12) is integrable over Ω_2 for functions such as $m(t,\theta) = \xi e^{-\theta t}$. In the appendix, we will establish that L is a continuous operator from the Hilbert space $L_2(\Omega_1)$ to $L_2(\Omega_2)$ as long as Ω_2 is a compact subset of Ω_1 which excludes zero and where $\Omega_1 = [0, \infty)$.

The choice of Ω_2 is not unique and, indeed, we arrived at our choice through both trial-and-error and our simulations. Therefore, at this point, it should be viewed as only a recommendation. More will be said about this in Section 5.1. Nonetheless, we observe that a solution to (12) will exist by the compactness of the parameter space, Θ, provided our objective function is continuous in θ for almost all $s \in \Omega_2$.

We do not solve problem (12) directly, since this could be a numerically intensive procedure, but rather, we investigate numerical approximations to the integral. In particular, in Section 5.1, we will be analyzing the numerical performance of two integral approximation schemes.

One will be the Riemann approximation to the integral in (12); that is, we solve

$$\min_{\theta \in \Theta} \sum_{j=1}^{k-1} [L[m(t,\theta)](s_j) - \sum_{i=1}^{n-1} e^{-s_j t_i} y_i \Delta t_i]^2 \Delta s_j, \tag{13}$$

where $\Delta s_j = s_{j+1} - s_j$. Again based on simulations, we have found that, if possible, a reasonable choice of the points s_j is to take $k = n$ and $s_j = t_j$ for all j provided $t_1 \neq 0$. If $t_1 = 0$, then we let $s_1 = \epsilon > 0$ and $s_j = t_j$ for $j = 2, ..., n$.

An alternative approach to solving (12) is through the use of integral Laplace approximations, which are closely related to saddlepoint approximations. We now describe such an approximation.

Laplace's method for integrals provides an approximation for integrals of the form,

$$\int_{\Omega_2} e^{nh(s,\theta)} ds, \tag{14}$$

for n large enough. According to the Laplace approximation, provided $h(s, \theta)$ has a unique maximizer, $\hat{s}_\theta$, for each fixed θ, the major contribution to the value of the integral arises from the immediate vicinity of those points of the interval of integration at which $e^{nh(s,\theta)}$ assumes its largest value for each θ [Erdelyi (1956)]. If a function $g(s, \theta)$ is non-vanishing for all s and each θ, then the choice of $h(s, \theta) = \frac{1}{n} \log g(s, \theta)$ is defined; thus, we see that if $g(s, \theta)$ is the integrand in (12), then $h(s, \theta)$ exists. We use the Laplace integral approximations to solve our least squares problem in the following way. First, fix a grid of parameters, θ's, that the experimenter suspects contains the true parameter. Next, for each θ in the grid, solve the following 1−dimensional problem; that is, find $\hat{s}_\theta$ such that

$$\max_{s \in \Omega_2} (g(s, \theta)) := g(\hat{s}_\theta, \theta), \tag{15}$$

where Ω_2 is the compact subset of Ω_1 given by (12) and $g(s, \theta)$ is the integrand in (12). Provided the objective function, $g(s, \theta)$, is a continuous function in s, a solution to problem (15) will exist.

Finally, for each θ in the grid, calculate either the *first order*, the *second order*, or the *third order* integral Laplace approximations (see the Appendix 7.2 for these details and definitions) and choose as the parameter estimate the θ which yields the minimum of the integral Laplace approximations.

The next theorem shows that under the assumption that Θ is compact and under standard regularity assumptions, the minimum of the approximations, as outlined in the previous paragraph, will be the minimum of the original problem as posed by (12).

Theorem 4.1 *If the parameter space is compact, if the objective function in (12) is continuous with respect to $\theta \in \Theta$, and if the integrand in (12) for each θ has a unique maximizer $\hat{s}_\theta \in \Omega_2$ in Ω_2, then the solution of problem (12) provided n is large enough agrees with the minimizer of the Laplace approximations.*

Proof. The proof is a standard continuity argument employing the fact that the objective function and the integral Laplace approximations are continuous functions on a compact space for each s. □

4.1 Large Sample Properties

In this subsection, we establish the large sample properties of the estimator of θ obtained from (12). In particular, using the theory of unbiased estimating equations it is straightforward to verify that the resulting estimator is consistent and asymptotically normal.

We can actually work with a more general form than the models in (7) or (8), of which those models are special cases. Instead of (7), write

$$\int_{\Omega_1} e^{-st} y(t) dG(t) = \int_{\Omega_1} e^{-st} m(t,\theta) dG(t) + \int_{\Omega_1} e^{-st} \varepsilon(t) dG(t)$$

where $G(t)$ can be thought of as the cdf of t. Now write

$$\begin{aligned} L[y](s) &= \int_{\Omega_1} e^{-st} y(t) dG(t), \\ L[m](s,\theta) &= \int_{\Omega_1} e^{-st} m(t,\theta) dG(t), \\ L[\varepsilon](s) &= \int_{\Omega_1} e^{-st} \varepsilon(t) dG(t), \end{aligned}$$

and note that $L[\varepsilon](s) \sim N(0, \Sigma(s))$, where

$$\Sigma(s) = E\left(\left[\int_{\Omega_1} e^{-st} \varepsilon(t) dG(t) \right] \left[\int_{\Omega_1} e^{-st} \varepsilon(t) dG(t) \right]' \right).$$

Now for a given weight function $W(s)$, a least squares fit will minimize

$$\int_{\Omega_s} [L[y](s) - L[m](s,\theta)]^2 dW(s),$$

and hence, provided that integrand is absolutely continuous with respect to θ and that the minimizer occurs in the interior of the parameter space, then it will be a solution to

$$\int_{\Omega_s} [L[y](s) - L[m](s,\theta)] \left[\frac{\partial}{\partial\theta} L[m](s,\theta)\right] dW(s) = 0. \tag{16}$$

Defining $\Psi_s(y,\theta) = [L[y](s) - L[m](s,\theta)]\left[\frac{\partial}{\partial\theta}L[m](s,\theta)\right]$, we see that (16) defines an estimating equation. As $E_\theta \Psi_s(Y,\theta) = 0$, (16) is an unbiased estimating equation. Following Carroll, Ruppert and Stefanski (1995, Appendix A.3) we have that the solution to (16), say $\hat{\theta}$, is a consistent estimator of θ. Moreover, as $n \to \infty$, a Taylor series argument will show that

$$\hat{\theta} - \theta \sim N\left(0, \Sigma_1^{-1}\Sigma_2\Sigma_1^{-1}\right),$$

where

$$\begin{aligned}
\Sigma_1 &= E_\theta\left[\int_{\Omega_1} -\frac{\partial^2}{\partial\theta^2} L[m](s,\theta) dW(s)\right] \\
\Sigma_2 &= E_\theta\left[\int_{\Omega_1} \Psi_s(Y,\theta) dW(s)\right]^2.
\end{aligned}$$

We now specialize to the case of (8) and (13), where we observe at times $t_1, \cdots, t_n$ and fit at points $s_1, \cdots, s_k$. If we consider $G(t)$ and $W(s)$ to be empirical distribution functions, we can write

$$\begin{aligned}
\Sigma_1 &= \frac{-1}{(n-1)(k-1)} \sum_{i=1}^{n-1}\sum_{j=1}^{k-1} e^{-s_j t_i} \frac{\partial^2}{\partial\theta^2} m(t_i,\theta)\Delta t_i \Delta s_j, \\
\Sigma_2 &= \frac{1}{[(n-1)(k-1)]^2} \sum_{i=1}^{n-1} \left\{ \left(\frac{\partial}{\partial\theta} m(t_i,\theta)\right)^2 \mathrm{var}(Y_i)\Delta t_i \right. \\
&\quad \left. \times \left[\sum_{j=1}^{k-1} e^{-4s_j t_i}\Delta s_j + 2\sum_{j'>j} e^{-2(s_j+s_{j'})t_i}\Delta s_j \Delta s_{j'}\right]\right\}.
\end{aligned}$$

5 Numerical Examples

We now look at the performance of our estimation procedure both in a simulation study and an example. One purpose of the simulation study is to

investigate the goodness of the Laplace/saddlepoint approximations to the objective functions. We also look at the Indomethacin absorption data set from [Kwan, et al (1975)] to see how our method will perform in practice.

5.1 Simulations

In this section we analyze the numerical performance of the Riemann approximation method (13) (referred to as (RA) in Table 5.1) and that of two Laplace integral approximation schemes–the second (2ND) and the third (3RD) order approximations (the first order approximation did not yield satisfactory results, and we do not report these results here).

We compare these results to those yielded by directly solving (12) (referred to as (INT)) and to those yielded by solving the standard least squares (SLS) problem. We do this over 500 simulated data sets on two models each with $m(t,\theta) = -\xi e^{-\theta t}$. Both of these models are a single compartment model representing, say, the single dose intravenous administration of a drug with the site of measurement being the same as the input site.

Model 1 has the true value of the parameter, θ, set to 50 with initial input $\xi = 100$. Model 2 has the true value of the parameter set to .1 and the initial input set to 5.

These results were obtained over a fixed grid consisting of 100 parameters that originated far from the true value and then clustered around the true parameter value. These models each have that $t_1 = 0$ (since the initial input is $m(0,\theta) = \xi$), thus for both we pick $s_1 = \epsilon = .00001$ following the discussion after (12). All other s_j are taken equaled to t_j on a uniformly spaced grid on the interval $[\epsilon, t_n]$ where $n = 30$ in one case and $n = 60$ in the other. For Model 1, t_n is .09, and for Model 2, t_n is 50.

In Table 5.1 we report the findings of our simulation study. The mean and variance estimates were calculated from the optimal parameter estimates obtained from analyzing each of the 500 data sets via each of the methods (SLS), (INT), (RA), (2ND) and (3RD). Overall, the standard least squares (SLS) yielded estimates which were closer in average value to the true parameter and with smaller variance. However, the results from (INT) and (RA) are comparable to those of (SLS) particularly for the case when $n = 60$. These results were obtained using the optimization and statistical analysis subroutines available from the IMSL numerical library.

Table 5.1

(500 Data Sets)

Model 1 (θ true = 50)

	$n = 60$		$n = 30$	
Method	Mean	Variance	Mean	Variance
(SLS)	49.99	.0732	50.00	.1458
(INT)	48.12	.2880	46.32	.3717
(RA)	48.12	.2880	46.32	.3709
(2ND)	48.70	43.72	48.98	38.51
(3RD)	48.30	48.61	48.74	41.84

Model 2 (θ true = .1)

	$n = 60$		$n = 30$	
Method	Mean	Variance	Mean	Variance
(SLS)	.1004	.0001	.1018	.0003
(INT)	.0889	.0001	.0789	.0001
(RA)	.0935	.0002	.0894	.0003
(2ND)	.1057	.0002	.1069	.0001
(3RD)	.1057	.00002	.1069	.0001

5.2 Indomethacin Data

In this section, using the same compartment model as in the simulation study, we fit the Indomethacin absorption data set from [Kwan, et al (1975)], where 6 healthy male volunteers each received an initial dose of indomethacin-2^{14}-C (about 4.16 $\mu g/ml$) by intravenous route. A total of 12 observations through time where made on each of the 6 subjects. Mean concentrations of the drug in the plasma are summarized in Figure 5.2; that is, each + corresponds to the average of the observations made at each time per subject. We plot this against the curves $m(t, \theta) = -4.16e^{-\theta t}$ with $\theta = 2.21$ (corresponds to dashed line in figure 5.2) and $\theta = 1.10$, and where t is in $\Omega_2 = [.0001, 8]$. These values of θ are the means reported on Table 5.2 for methods (SLS) and (RA), respectively. Table 5.2 is similar to Table 5.1 except that now our numerical entries were calculated from the data gathered for each of the 6 individuals in the study. We only report the findings using methods (SLS), (INT), and (RA) and not those using methods (2ND) and (3RD) since they did not yield satisfactory results when analyzing this particular data set.

Table 5.2

(6 Data Sets, $n = 12$)

Method	Mean	Variance
(SLS)	2.215	.2482
(INT)	1.080	.0225
(RA)	1.101	.0337

6 Summary and Discussion

We presented an alternative method for statistical parameter estimation based on integral transform techniques and integral approximations schemes. This method appears to be promising for compartment models, especially those with many compartments. We say this, because our method yields a closed form expression of the concentration curve at the measurement site (even though it is in terms of the variable of the transformation) without implicitly necessitating that we solve the system of differential equations– that is, form the spectral decomposition of A_θ which need not exist. Clearly, both theory and simulations are needed in order to establish the algorithm's performance or superiority on small or medium size data sets over existing methods of estimation, namely standard least squares. We also note that our method could, no doubt, be made more efficient by making it more adaptive to a data set, perhaps by using a variable rather than a fixed grid.

It would also be interesting to see if the pattern that we have witnessed in the simulation study and the analysis of the real data set, between the estimates obtained from solving (12) and those obtained from solving its Riemann approximation, (13), persists for a many compartment model.

We end with a word of caution, in that estimation in the transformed space brings along with it a degree of arbitrariness. Although our estimates have performed well in small sample studies, and have reasonable large sample properties, the choice of points s_j is not yet well established. In Section 5.1 we arrived at the choice of s_j by experimentation, and at this stage we are not prepared to make a general recommendation which would be certainly desirable, especially in light of the simulation study and of Theorem 7.1.

7 Appendix

The appendix contains technical details of the estimation method, addressing properties of the transformation and the integral approximation.

7.1 Continuity of the Laplace Transform

Theorem 7.1 *Let L be as given by (9), $\Omega_1 = [0, \infty)$ and Ω_2 a compact subset of Ω_1 that excludes zero. Suppose $h \in L_2(\Omega_1)$ is such that, $L[h](s)$ exists; then L is a continuous operator from the Hilbert spaces $L_2(\Omega_1)$ to $L_2(\Omega_2)$.*

Proof. The theorem will follow from a direct application of Cauchy-Schwarz's inequality and Fubini's theorem. We first re-write (9), apply Cauchy-Schwarz's inequality and use the fact that $|e^{-st}| \leq 1$ since both s and t are non negative to obtain the following.

$$\begin{aligned} (\int_0^\infty e^{-st/2} e^{-st/2} h(t)dt)^2 &\leq (\int_0^\infty e^{-st}dt)(\int_0^\infty e^{-st}h^2(t)dt) \\ &\leq \frac{1}{s}\int_0^\infty h^2(t)dt. \end{aligned} \tag{17}$$

Therefore, by (17) and Fubini's theorem we have that

$$\begin{aligned} ||L[h](s)||_{L_2(\Omega_2)} &\leq \int_0^\infty (\int_{\Omega_2} \frac{1}{s} ds) h^2(t)dt \\ &\leq k||h(t)||^2_{L_2(\Omega_1)}, \quad \text{for some } k > 0. \end{aligned}$$

Thus, the theorem is established. □

7.2 Laplace Approximations

The integral Laplace approximation scheme mentioned in Section 4 can be obtained as follows.

We do a Taylor series expansion of $h(s, \theta)$ about its maximizer, $\hat{s}_\theta$, but we work with the first three terms of the series, that is,

$$h(s, \theta) \simeq h(\hat{s}_\theta, \theta) + \frac{(s - \hat{s}_\theta)^2}{2!} h''(\hat{s}_\theta, \theta) + \frac{(s - \hat{s}_\theta)^3}{3!} h'''(\hat{s}_\theta, \theta)$$

where $h''(\hat{s}_\theta,\theta) = \frac{\partial^2 h(s,\theta)}{\partial s^2}\Big|_{s=\hat{s}_\theta}$ and $h'''(\hat{s}_\theta,\theta) = \frac{\partial^3 h(s,\theta)}{\partial s^3}\Big|_{s=\hat{s}_\theta}$. Then the following approximation, which is valid within a neighborhood of $\hat{s}_\theta$, yields

$$\int_{\Omega_2} e^{nh(s,\theta)}ds \simeq e^{nh(\hat{s}_\theta,\theta)}[\int_{\Omega_2} e^{n\frac{(s-\hat{s}_\theta)^2}{2!}h''(\hat{s}_\theta,\theta)}e^{n\frac{(s-\hat{s}_\theta)^3}{3!}h'''(\hat{s}_\theta,\theta)}ds].$$

Expanding the exponential cubic term into its series expansion, we obtain the more refined result

$$\begin{aligned}\int_{\Omega_2} e^{nh(s,\theta)}ds \;\simeq\; & \int_{\Omega_2} e^{n\frac{(s-\hat{s}_\theta)^2}{2!}h''(\hat{s}_\theta,\theta)} \\ & \times\left[1+n\frac{(s-\hat{s}_\theta)^3}{3!}h'''(\hat{s}_\theta,\theta)+n^2\frac{(s-\hat{s}_\theta)^6}{2!(3!)^2}[h'''(\hat{s}_\theta,\theta)]^2+R_n\right]ds,\end{aligned} \tag{18}$$

where R_n is the remainder term. We call the integral approximations in (18) the *first order* approximation if it excludes the last three terms in the right-hand side (including R_n). The *second order* approximation if it excludes the last two terms; lastly the *third order* approximation includes all the terms except R_n. More precisely, letting $\Phi(\cdot)$ denote the standard normal cdf, and taking $\Omega_2 = [t_1, t_n]$, the first order approximation is

$$\begin{aligned}\int_{t_1}^{t_n} e^{nh(s,\theta)}ds \;\simeq\; & e^{nh(\hat{s}_\theta,\theta)}\sqrt{\frac{2\pi}{nh''(\hat{s}_\theta,\theta)}} \\ & \times\left\{\Phi[\sqrt{nh''(\hat{s}_\theta,\theta)}(t_n-\hat{s}_\theta)]-\Phi[\sqrt{nh''(\hat{s}_\theta,\theta)}(t_1-\hat{s}_\theta)]\right\}.\end{aligned}$$

Using the facts that

$$\begin{aligned}\int y^3e^{-ay^2/2}dy &= \frac{-1}{2a}\left[y^2+\frac{1}{a}\right]e^{-ay^2/2} \\ \int y^6e^{-ay^2/2}dy &= \frac{-1}{2a}\left[y^5+\frac{5y^3}{2a}+\frac{15y}{4a}\right]e^{-ay^2/2}+30\sqrt{\frac{\pi}{a^7}}\Phi(\sqrt{2a}y)\end{aligned}$$

similar expressions can be derived for the second and third order approximations.

References

[1] Bagchi U., Ord J.K. and Sullivan R.S. (1986) Parameter estimation for convolutions and compound distributions, *Operations Research Letters*, Vol. 4, No. 6, pp. 301-308.

[2] Carroll, R. J. Ruppert, D. and Stefanski, L. (1995). *Measurement Error in Nonlinear Models.* London: Chapman and Hall

[3] Erdelyi A. (1956) Asymptotic Expansions, *Dover Publications, Inc.,*, New York, NY.

[4] Feigin P.D., Tweedie R.L. and Belyea C. (1983) Weighted area techniques for explicit parameter estimation in multi-stage models, *Austral. J. Statist.*, Vol. 25, No. 1, pp. 1-16.

[5] Feuerverger A., McDunnough P. (1981a) On some Fourier methods of inference, *Journal of the American Statistical Association*, Vol. 76, No. 374, pp. 379-387.

[6] Feuerverger A., McDunnough P. (1981b) On the efficiency of empirical characteristic function procedures, *J. R. Statist. Soc.* **B**, Vol. 43, No. 1, pp. 20-27.

[7] Feuerverger A., McDunnough P. (1984) On statistical transform methods and their efficiency, *The Canadian Journal of Statistics*, Vol. 12, No. 4, pp. 303-317.

[8] Kwan K.C., Breault G.O, Umbenhauer E.R., McMahon F.G., and Duggan D.E. (1975) Kinetics of Indomethacin Absorption, Elimination, and Enterohepatic Circulation in Man, *Journal of Pharmacokinetics and Biopharmaceutics,* Vol. 4, pp. 255-280.

[9] Laurence A.F. and B.J.T. Morgan (1987) Selection of the transformation variable in the Laplace transform method of estimation, *Austral. J. Statist.*, Vol. 29, No. 2, pp. 113-127.

[10] Moler C. and Van Loan C. (1978) Nineteen dubious ways to compute the exponential of a matrix, *SIAM Rev.,* Vol. 20, pp. 801-836.

[11] Schuh H.J. and Tweedie R.L. (1979) Parameter estimation using transform estimation in time-evolving models, *Mathematical Biosciences* Vol. 45, pp. 37-67.

[12] Walter, G.G. (1979) Compartmental models, digraphs, and Markov chains, *Compartmental Analysis of Ecosystem Models*, Statistical Ecology Series, J.H. Matis, B.C. Patten, and G.C. White (eds), Vol. 10, pp. 295-310.

[13] Wells C.E. and Bryant J.L. (1984) An adaptive estimation procedure using the Laplace transformation, *IIE Transactions*, Vol. 17, No. 3, pp. 242-251.

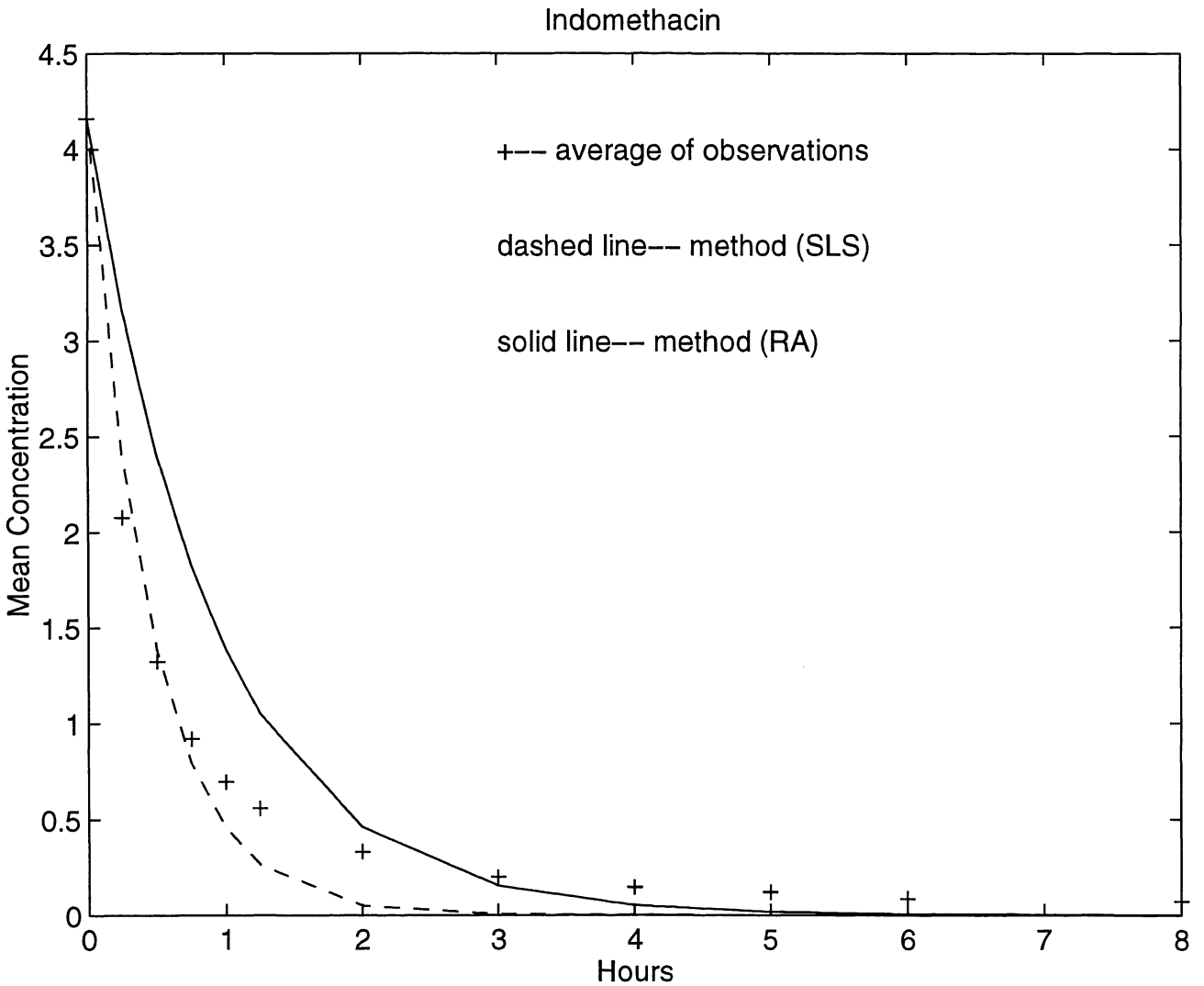

Figure 1: Indomethacin Absorption Data